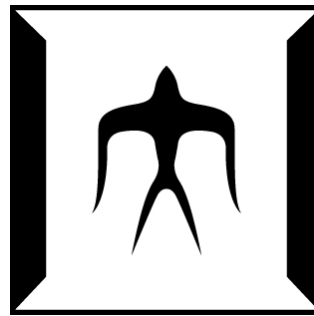


「山吹」をめぐる和歌語彙の空間



山元啓史

東京工業大学大学院社会理工学研究科

カリフォルニア大学サンディエゴ校環太平洋大学院大学

December 10th 2011

本論の主旨（１）

歌ことばの特徴は何か？

専門家に聞いてみる

「 の特徴は何か？ 」

答えられない or

歌ことば辞典に記述なし

本論の主旨（２）

専門家が指摘しない語の意味、
その理由は？

1. そういうものは重要でない。
2. 常識であって記述するほどのものではない。
3. そういうものを本当に知らない。
4. **言われればそうだが、思いつかなかった。**

本論の主旨（3）

専門家でも指摘できない語

計算手続きで指摘する

辞書開発に役立てる

（関係の記述）

「山吹」をめぐって

八代集での「山吹」の出現頻度

勅撰集名	頻度
古今集	6
後撰集	4
拾遺集	6
後拾遺集	6
金葉集	5
詞花集	3
千載集	7
新古今集	7
計	44

厚見王、万葉集 1435

蝦鳴く

甘南備河に

かげ見えて

今か咲くらむ

山吹の花

厚見王、万葉集 1435

蝦鳴く

甘南備河に

かけ見えて

今か咲くらむ

山吹の花

芭蕉

古池や

蛙飛びこむ

水の音

芭蕉

山吹や

蛙飛びこむ

水の音

歌川広重「山吹に蛙」

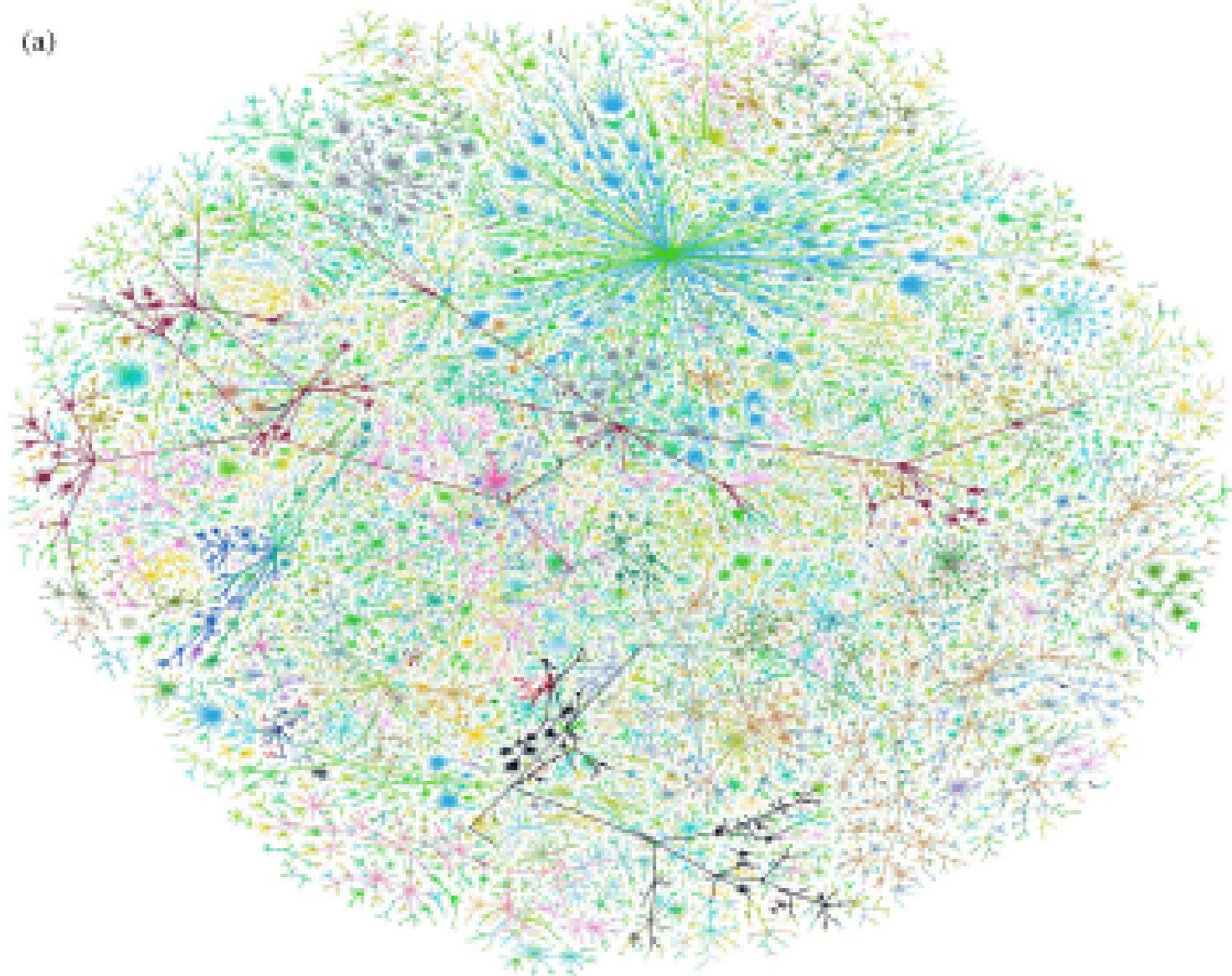


本論の目的

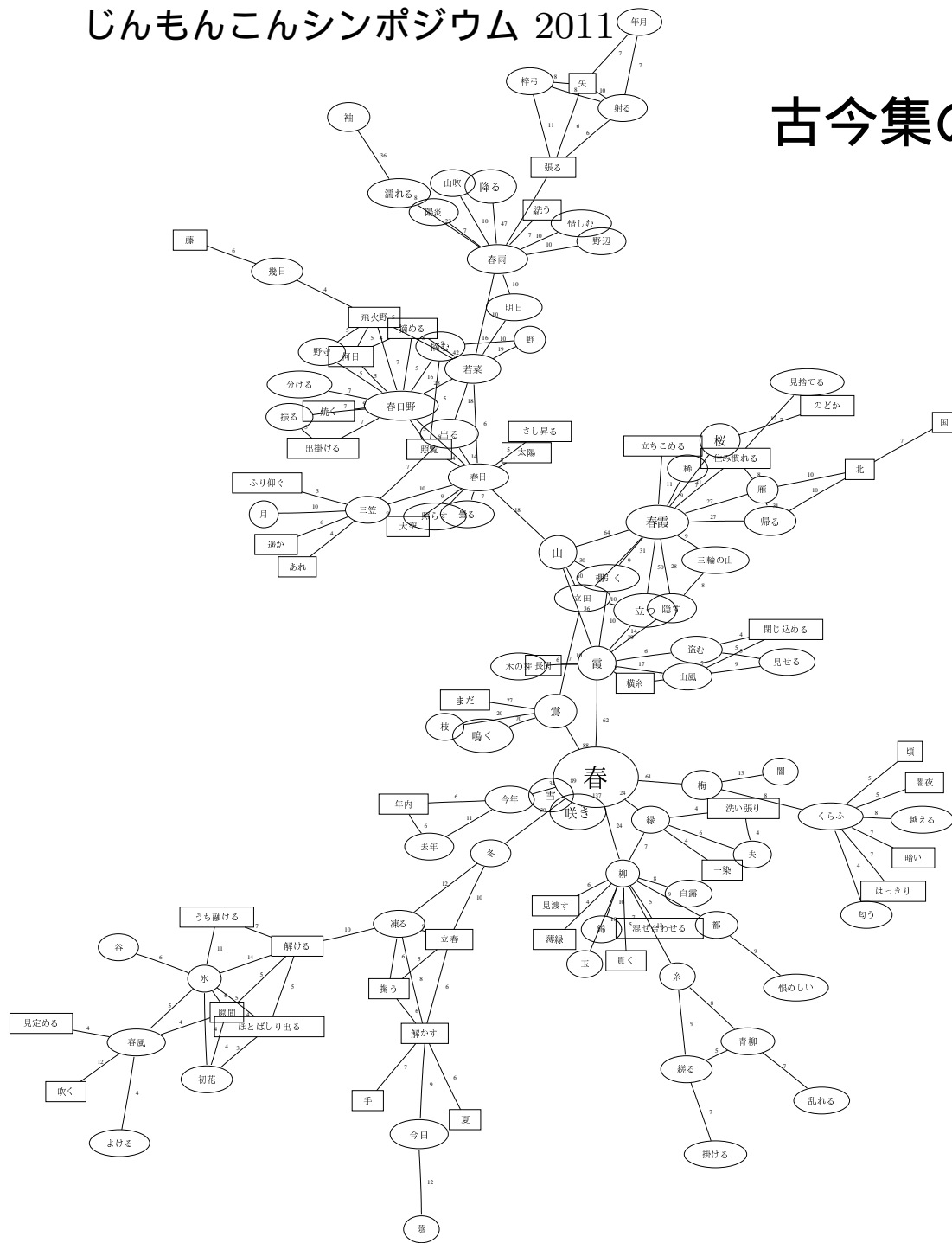
計算手続きによる
「山吹」をめぐる
語彙の空間の検討

インターネットのネットワーク構造

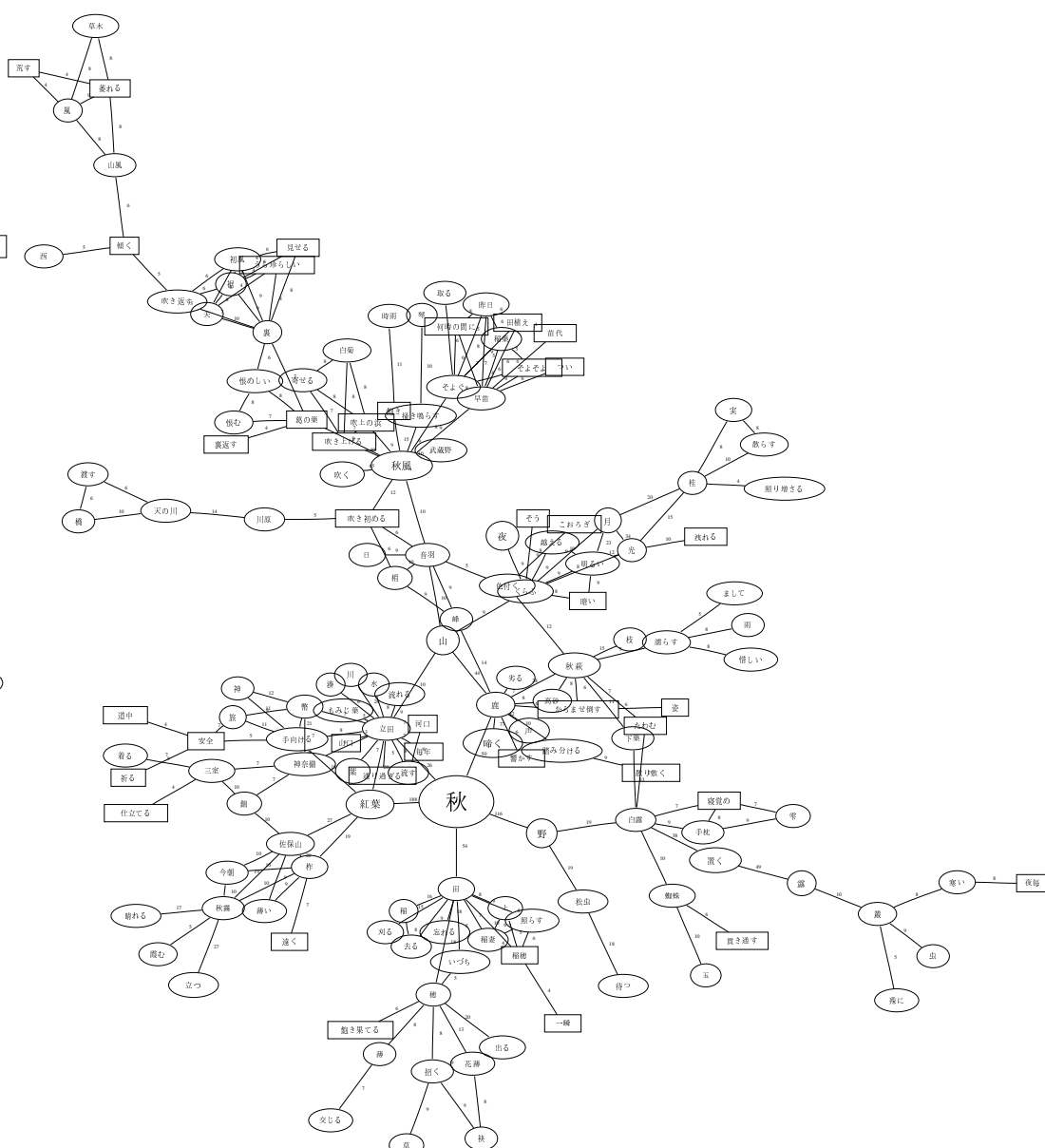
(a)



古今集の春歌と秋歌の空間



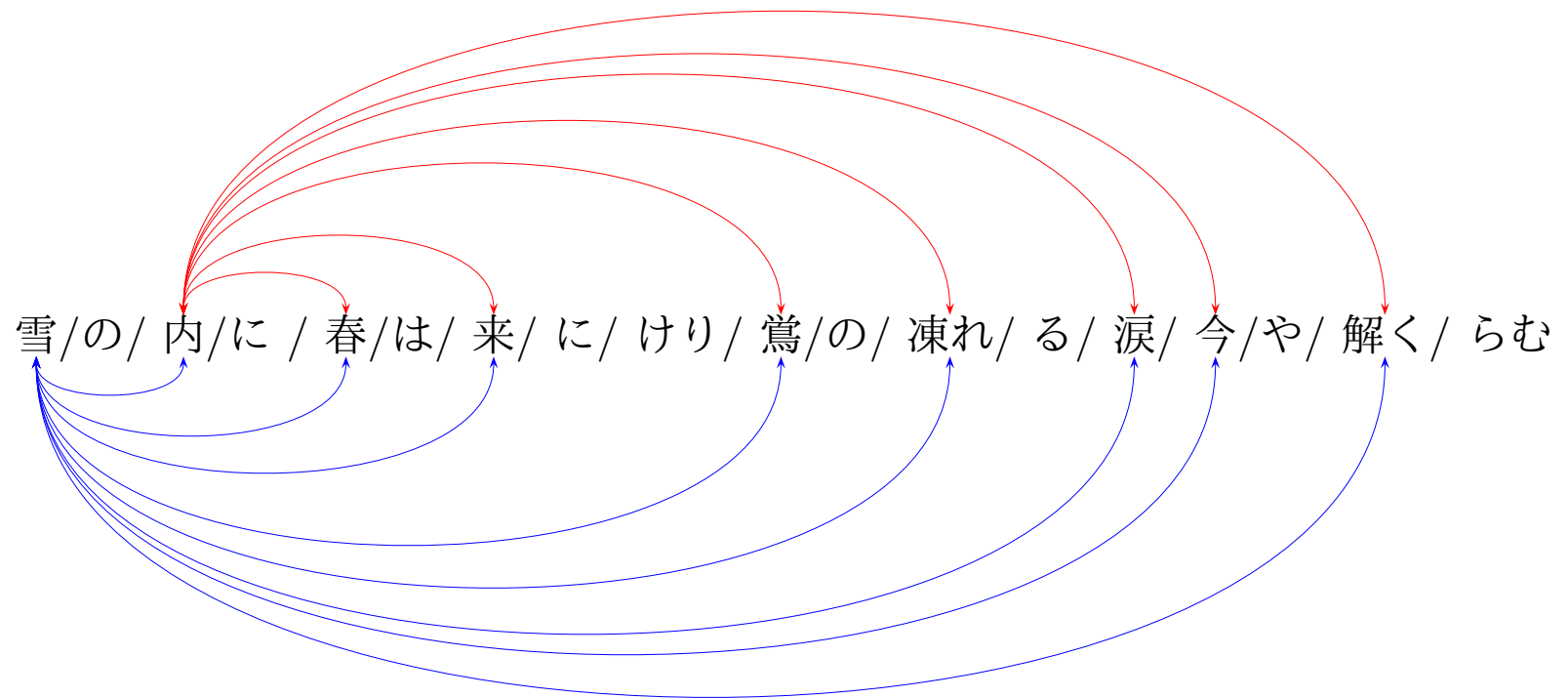
春 (90/652.2.68): CT cw.>15.6; non-dist=off: idf=on(2)



秋 (133/988.2.26): CT cw.>16; non-dist=off: idf=on(2)

グラフによる分析

- 共出現パターンとグラフ理論



- 共出現パターンの重み付け計算

グラフによる分析

- 共出現パターンとグラフ理論
- 共出現パターンの重み付け計算

グラフによる分析

- 共出現パターンとグラフ理論
- 共出現パターンの重み付け計算

$$w(t, d) = (1 + \log tf(t, d)) \cdot idf(t) \quad (1)$$

$$cw(t_1, t_2, d) = (1 + \log ctf(t_1, t_2, d)) \cdot cidf(t_1, t_2) \quad (2)$$

$$cidf(t_1, t_2) = \sqrt{idf(t_1) \cdot idf(t_2)} \quad (3)$$

$$idf(t) = \log \frac{N}{df(t)} \quad (4)$$

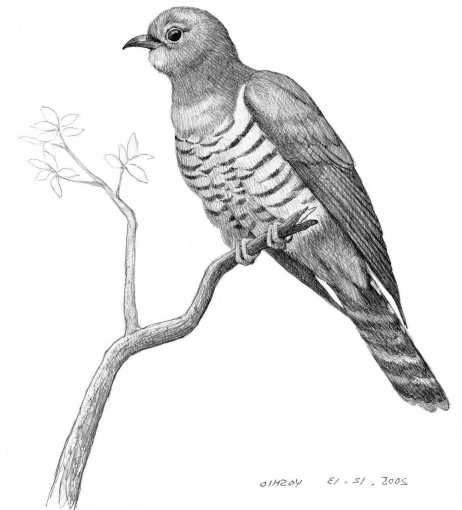
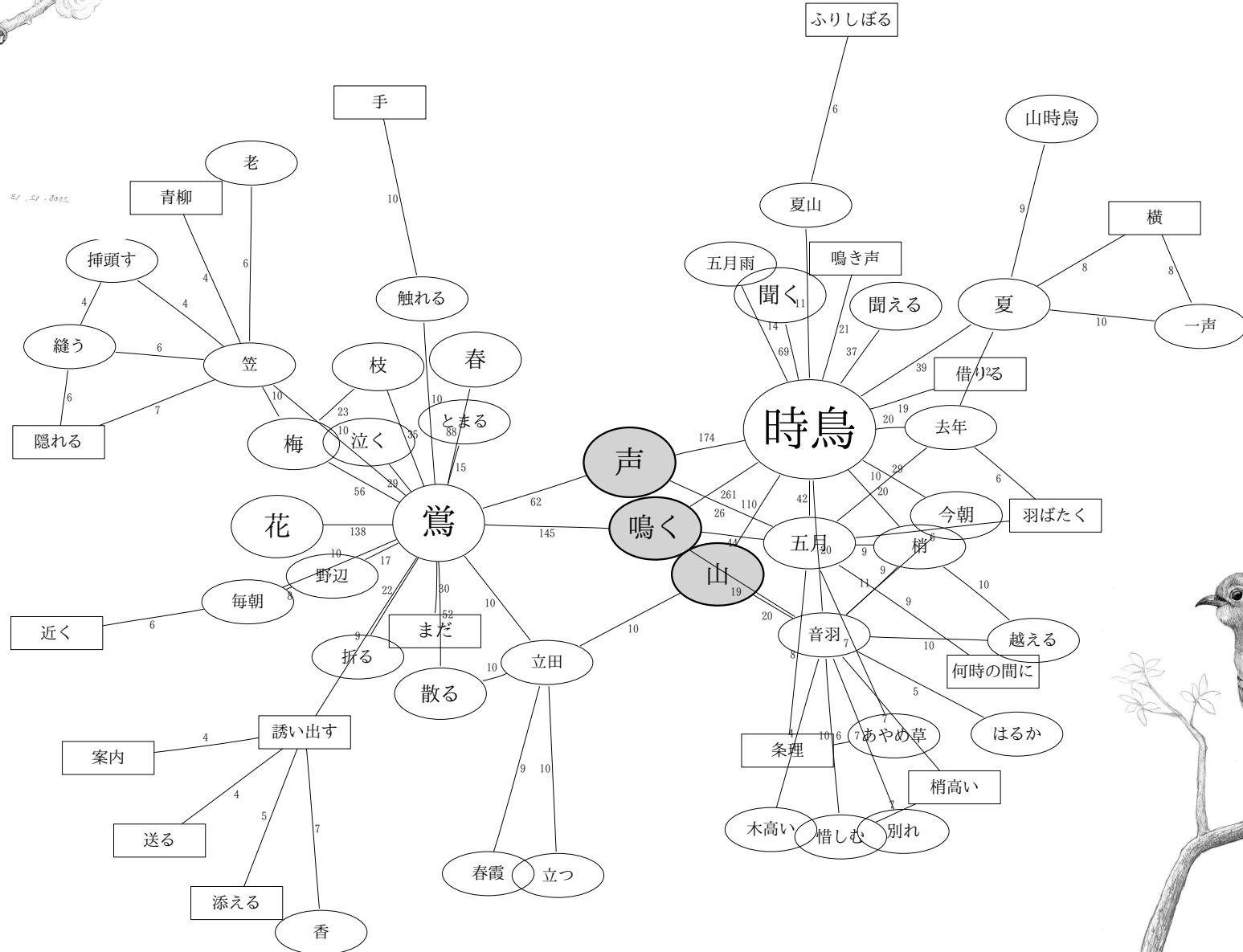
(1)...Manning, 1999

(4)...Spärck Jones, 1972

こういうモデルを作って分析する



©IMZDY 2011. 5/11. 2005



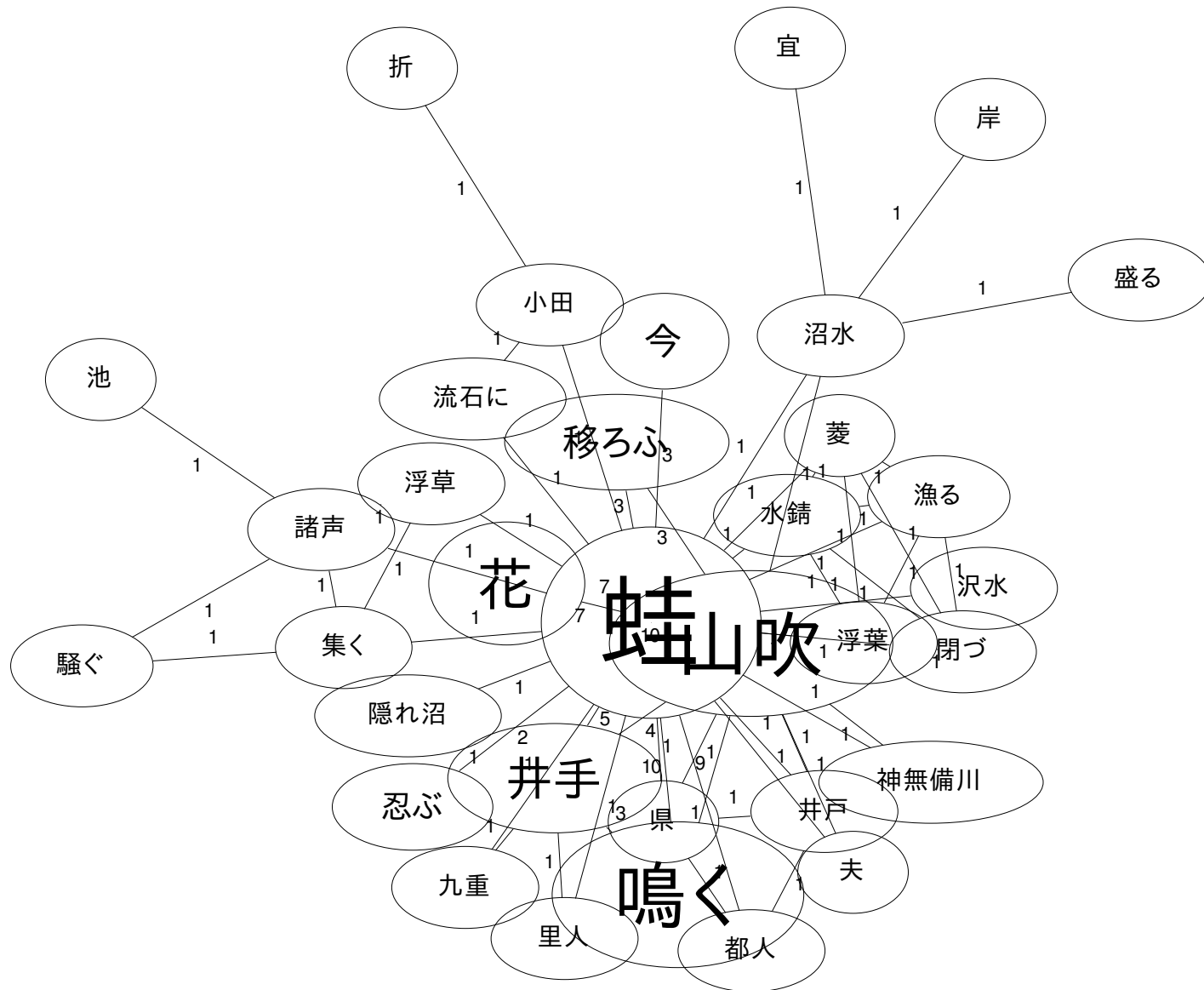
©IMZDY 2011. 5/11. 2005

方法：材料

- 国文学研究資料館開発正保版本「二十一代集」
- 長歌を除く 9484 首の和歌テキスト
(シソーラスの作成はすべての和歌に対して)
- kh で単位分割 (短単位) し、
- 異形同語 (立田 / 竜田 / 龍田) の問題 t2c でシソーラスコードをつける。
- 八代集シソーラスの開発 分類語彙表を利用
- 一般語 (48732)、地名 (1408)、人名 (49)

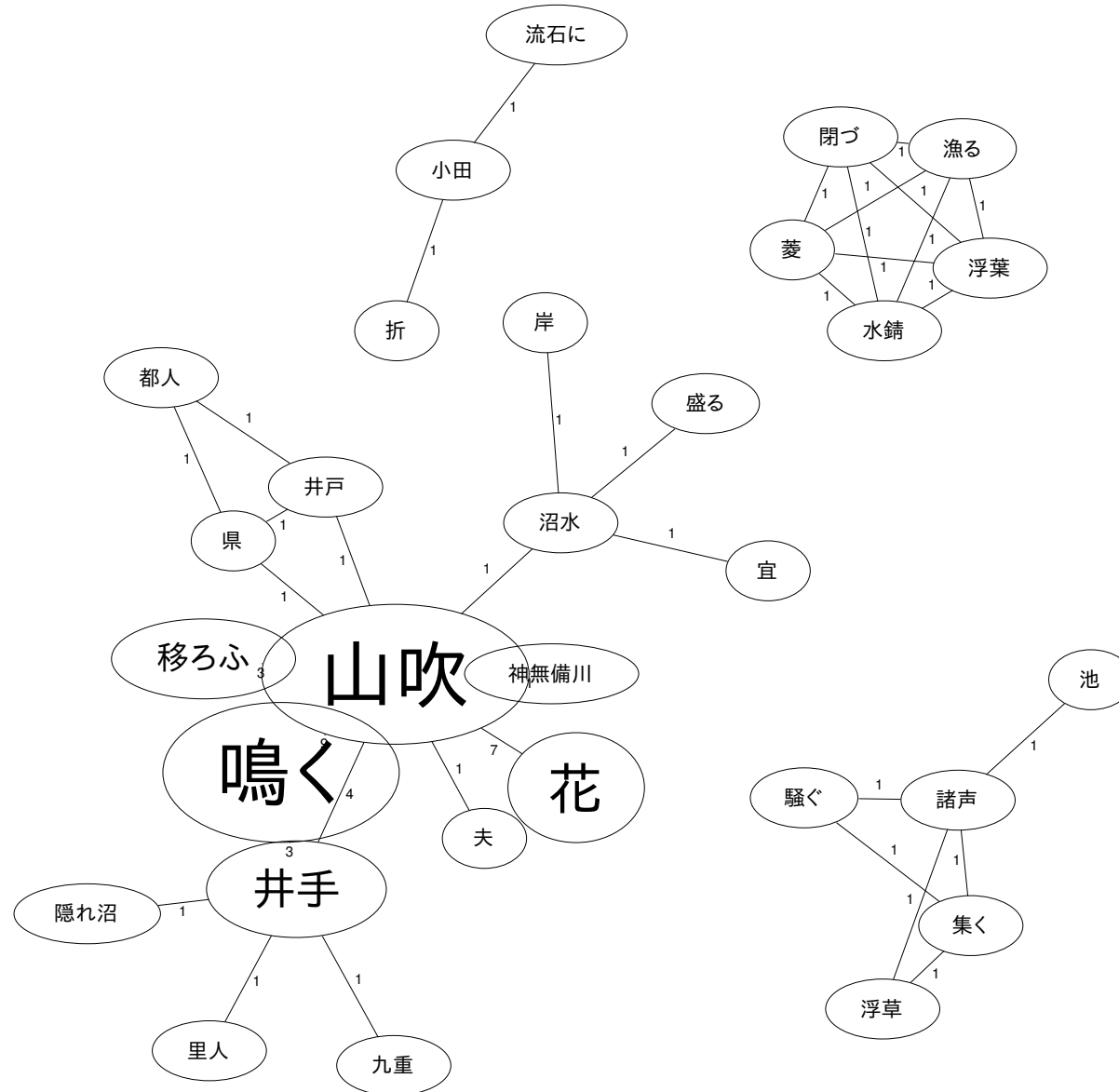
結果

蛙から 見てみる



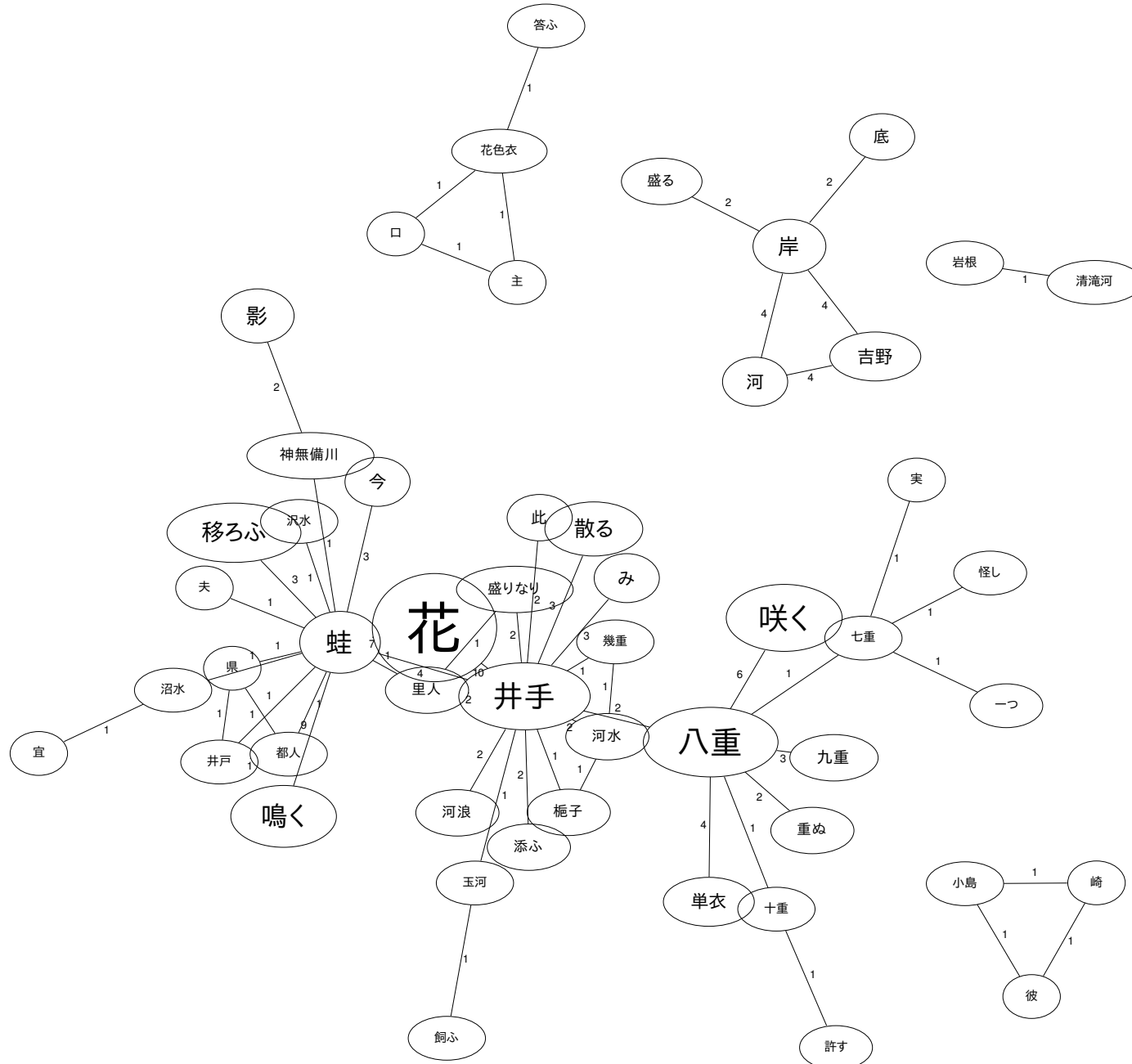
蛙 (15/15/15, 6.45) cw > 2.50 K:1-8 U:1 L:0.00 M:7 Z:1.00

蛙から
見てみる
pruning



山吹から
見てみる
pruning

蛙
井手
八重
...HUB



「八重」と「山吹」の関係

- 「八重」を検索してみる
 - 「八重（やへ）」「八重葎（やへむぐら）」
 - 「八重桜（やへざくら）」「八重雲（やへくも）」
 - 「八重菊（やへぎく）」
- 「八重山吹」
 - － 「山吹」の花には2種類：一重のものと八重のもの
 - － この解説は歌ことば関連の辞典には見られない

「八重山吹」の実態

- モデルから探してみる。

八重さける
かひこそなけれ
山吹の
ちらはひとへも
あらしとおもへは

詞花集 46 番 読人不知

おわりに

- 「山吹」とともに詠まれる語彙の検討
- 辞典類: 「山吹」と「八重」の関連性
述べられていない。
- モデル: 「山吹」と「蛙」「井手」「八重」の関係
示された。そして「**八重**山吹」であることも...
- 外在的基準の利用
- 辞典などの見出し語の検討
- モデルによる言語記述の有効性